(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 352 959 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
15.10.2003 Bulletin 2003/42

(51) Int Cl.[7]: **C12N 15/10**, C07K 16/00, C12Q 1/68

(21) Application number: **03076606.7**

(22) Date of filing: **26.01.1998**

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **24.01.1997 GB 9701425**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
98901380.0 / 0 988 378

(71) Applicant: **Bioinvent International AB**
223 70 Lund (SE)

(72) Inventors:
- **Borrebaeck, Carl Arne Krister**
  245 62 Hjärup (SE)
- **Söderlind, Ulf Hans Eskil**
  247 35 Söndra Sandby (SE)

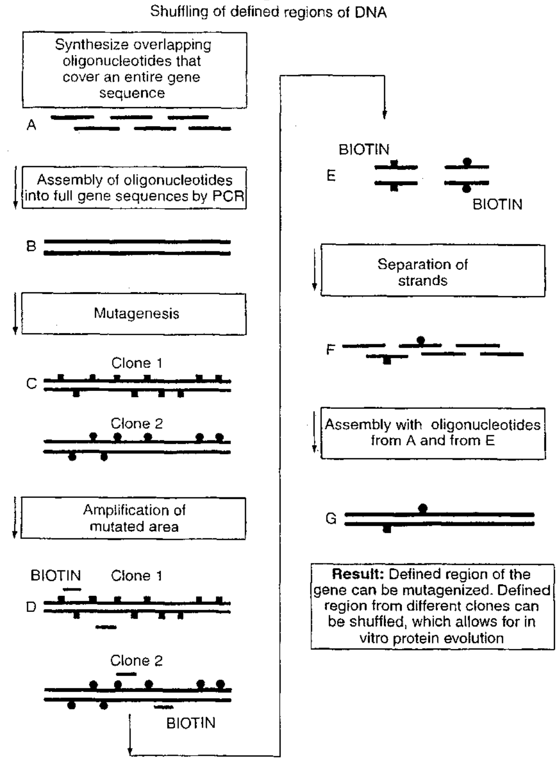
(74) Representative: **Thomas, Philip John Duval et al**
Eric Potter Clarkson,
Park View House,
58 The Ropewalk
Nottingham NG1 5DD (GB)

Remarks:
This application was filed on 26 - 05 - 2003 as a divisional application to the application mentioned under INID code 62.

(54) A method for in vitro molecular evolution of protein function

(57) The present invention relates to a method for in vitro creation of molecular libraries evolution of protein function. Particularly, it relates to variability and modification of protein function by shuffling polynucleotide sequence segments. A protein of desired characteristics can be obtained by incorporating variant peptide regions (variant motifs) into defined peptide regions (scaffold sequence). The variant motifs can be obtained from parent DNA which has been subjected to mutagenesis to create a plurality of differently mutated derivatives thereof or they can be obtained from in vivo sequences. These variant motifs can then be incorporated into a scaffold sequence and the resulting coded protein screened for desired characteristics. This method is ideally used for obtaining antibodies with desired characteristics by isolating individual CDR DNA sequences and incorporating them into a scaffold which may, for example, be from a totally different antibody.

Fig.1.
Shuffling of defined regions of DNA



EP 1 352 959 A1

Description

Field of the invention

[0001] The present invention relates to a method for in vitro molecular evolution of protein function. Particularly, but not exclusively, it relates to the shuffling of polynucleotide sequence segments within a coding sequence.

Background of the invention

[0002] Protein function can be modified and improved in vitro by a variety of methods, including site directed mutagenesis (Moore et al, 1987) combinatorial cloning (Huse et al, 1989; Marks et al, 1992) and random mutagenesis combined with appropriate selection systems (Barbas et al, 1992).

[0003] The method of random mutagenesis together with selection has been used in a number of cases to improve protein function and two different strategies exist. Firstly, randomisation of the entire gene sequence in combination with the selection of a variant (mutant) protein with the desired characteristics, followed by a new round of random mutagenesis and selection. This method can then be repeated until a protein variant is found which is considered optimal (Moore et al, 1996). Here, the traditional route to introduce mutations is by error prone PCR (Leung et al, 1989) with a mutation rate of ≈0.7%.

[0004] Secondly, defined regions of the gene can be mutagenized with degenerate primers, which allows for mutation rates up to 100% (Griffiths et al, 1994; Yang et al, 1995). The higher the mutation rate used, the more limited the region of the gene that can be subjected to mutations.

[0005] Random mutation has been used extensively in the field of antibody engineering. In vivo formed antibody genes can be cloned in vitro (Larrick et al, 1989) and random combinations of the genes encoding the variable heavy and light genes can be subjected to selection (Marks et al, 1992). Functional antibody fragments selected can be further improved using random mutagenesis and additional rounds of selections (Hoogenboom et al, 1992).

[0006] The strategy of random mutagenesis is followed by selection. Variants with interesting characteristics can be selected and the mutagenized DNA regions from different variants, each with interesting characteristics, are combined into one coding sequence (Yang et al, 1995). This is a multi-step sequential process, and potential synergistic effects of different mutations in different regions can be lost, since they are not subjected to selection in combination. Thus, these two strategies do not include simultaneous mutagenesis of defined regions and selection of a combination of these regions. Another process involves combinatorial pairing of genes which can be used to improve e.g. antibody affinity (Marks et al, 1992). Here, the three CDR-regions in each variable gene are fixed and this technology does not allow for shuffling of individual CDR regions between clones.

[0007] Selection of functional proteins from molecular libraries has been revolutionized by the development of the phage display technology (Parmley et al, 1987; McCafferty et al, 1990; Barbas et al, 1991). Here, the phenotype (protein) is directly linked to its corresponding genotype (DNA) and this allows for directly cloning of the genetic material which can then be subjected to further modifications in order to improve protein function. Phage display has been used to clone functional binders from a variety of molecular libraries with up to 10^{11} transformants in size (Griffiths et al, 1994). Thus, phage display can be used to directly clone functional binders from molecular libraries, and can also be used to improve further the clones originally selected.

[0008] Random combination of DNA from different mutated clones is a more efficient way to search through sequence space. The concept of DNA shuffling (Stemmer, 1994) utilises random fragmentation of DNA and assembly of fragments into a functional coding sequence. In this process it is possible to introduce chemically synthesised DNA sequences and in this way target variation to defined places in the gene which DNA sequence is known (Crameri et al, 1995). In theory, it is also possible to shuffle DNA between any clones. However, if the resulting shuffled gene is to be functional with respect to expression and activity, the clones to be shuffled have to be related or even identical with the exception of a low level of random mutations. DNA shuffling between genetically different clones will generally produce non-functional genes.

Summary of the invention

[0009] At its most general the present invention provides a method of obtaining a polynucleotide sequence encoding a protein of desired characteristics comprising the steps of incorporating at least one variant nucleotide region (variant motif) into defined nucleotide regions (scaffold sequence) derived from a parent polynucleotide sequence. The new assembled polynucleotide sequence may then be expressed and the resulting protein screened to determine its characteristics.

[0010] The present method allows protein characteristics to be altered by modifying the polynucleotide sequence encoding the protein in a specific manner. This may be achieved by either a) replacing a specified region of the nucleotide sequence with a different nucleotide sequence or b) by mutating the specified region so as to alter the nucleotide sequence. These specified regions (variant motifs) are incorporated within scaffold or framework regions (scaffold sequence) of the original polynucleotide sequence (parent polynucleotide sequence) which when reassembled will encoded a protein of altered characteristics. The characteristics of the encoded protein are altered as a result of the amino acid sequence being changed corresponding to the changes in the coding polynucleotide sequence.

[0011] Rather than modifying a sequence at random and then relying on extensive screening for the desired coded protein, the present inventors have found it desirable to provide a method which modifies selected segments (variant motifs) of a protein while maintaining others.

[0012] The variant motifs may be segments of nucleotide sequence that encode specified regions of a protein. For example, functional regions of a protein (e.g. loops) or CDR regions in an antibody.

[0013] The scaffold sequence may be segments of nucleotide sequence which it is desirable to maintain, for example they may encode more structural regions of the protein, e.g. framework regions in an antibody.

[0014] The variant motifs may be nucleotide segments which originated from the same polynucleotide sequence as the scaffold sequence, i.e. the parent polynucleotide sequence, but which have been mutated so as to alter the coding sequence from that in the parent. For example, the parent polynucleotide sequence may encode an antibody. The nucleotide sequences encoding the CDR regions of the antibody (variant motifs) may be selected from the remaining coding sequence of the parent polynucleotide, mutated and then reassembled with scaffold sequence derived from the remaining coding sequence. The expressed antibody will differ from the wild type antibody expressed by the parent polynucleotide in the CDR regions only.

[0015] Alternatively, the variant motif may be derived from a polynucleotide sequence encoding a protein sequentially related to the protein encoded by the parent polynucleotide sequence. For example, the CDR regions from one antibody (antibody A) may be replaced by the CDR regions of another antibody (antibody B).

[0016] In each case the resulting expressed protein can be screened for desired characteristics. Desirable characteristics may be changes in the biological properties of the protein. For example, the tertiary structure of the protein may be altered. This may affect its binding properties, the ability for it to be secreted from cells or into cells or, for enzymes, its catalytic properties. If the protein is an antibody or part thereof it may be desirable to alter its ability to specifically bind to an antigen or to improve its binding properties in comparison to the parent antibody.

[0017] According to one aspect of the present invention, there is provided a method of obtaining a protein of desired characteristics by incorporating variant peptide regions (variant motifs) into defined peptide regions (scaffold sequence), which method comprises the steps of:

(a) subjecting parent polynucleotide sequence encoding one or more protein motifs to mutagenesis to create a plurality of differently mutated derivatives thereof, or obtaining parent polynucleotide encoding a plurality of variant protein motifs of unknown sequence,
(b) providing a plurality of pairs of oligonucleotides, each pair representing spaced-apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif,and using each said pair of oligonucleotides as amplification primers to amplify the intervening motif;
(c) obtaining single-stranded nucleotide sequence from the thus-isolated amplified nucleotide sequence; and
(d) assembling nucleotide sequence encoding a protein by incorporating nucleotide sequences derived from step (c) above with nucleotide sequence encoding scaffold sequence.

[0018] The method may further comprise the step of expressing the resulting protein encoded by the assembled nucleotide sequence and screening for desired properties.

[0019] Preferably the parent polynucleotide sequence is DNA from which is derived DNA sequences encoding the variant motifs and scaffold sequences.

[0020] Preferably the pairs of oligonucleotides are single-stranded oligonucleotide primers. One of said pair may be linked to a member of a specific binding pair (MSBP). The MSBP is preferably biotin, whose specific binding partner could for example be streptavidin. By using the specific binding pair the amplified nucleotide sequences may be isolated.

[0021] Random mutation can be accomplished by any conventional method; but a suitable method is error-prone PCR.

[0022] The protein in question could, for example, be an antibody or antibody fragment having desirable characteristics. Example of antibody fragments, capable of binding an antigen or other binding partner, are the Fab fragment consisting of the VL, VH, C1 and CH1 domains; the Fd fragment consisting of the VH, and CHI1 domains; the Fv fragment consisting of the VL and VH domains of a single arm of an antibody; the dAb fragment which consists of a VH domain; isolated CDR regions and F(ab')2 fragments, a bivalent fragment including two Fab fragments linked by a disulphide bridge at the hinge region. Single chain Fv fragments are also included.

[0023] In one approach, after randomly mutating DNA encoding the antibody, or a portion of that DNA (eg that which encodes the Fab regions or variable regions), oligonucleotide primers could be synthesised corresponding to sequences bounding the CDRs (the variant motifs), so that DNA encoding the CDRs are amplified, along with any mutations that may have occurred in the CDRs. These can be incorporated in the reassembly of the antibody coding sequence, using the amplified CDR DNA sequences and the unmutated scaffold framework (FR) DNA sequences, resulting in the expression of an antibody which has a novel combination of CDRs, and potentially having altered properties which can be selected or screened for in conventional manner.

[0024] In another approach, rather than mutate CDRs and reassembling them back into an antibody which will be closely related to the parent antibody from which the CDRs were derived, the CDRs may be taken from one or more existing antibodies, but be of unknown sequence. Using oligonucleotide primers representing sequences bounding the various CDRs, the individual CDRs can be amplified, isolated and assembled into a predetermined scaffold.

[0025] Of course, combinations of the foregoing approaches could be used, with CDRs taken from one or more parent antibodies, and assembled into a scaffold to produce a completely new, secondary antibody, then, after screening to obtain a secondary antibody with desired characteristics, the DNA encoding it could be mutated, the CDRs amplified and isolated, and then reassembled with unmutated non-CDR (scaffold) DNA from the secondary antibody, to produce variants of the secondary antibody which are mutated in the CDRs, and which can be screened for improved properties with respect to the originally selected secondary antibody.

[0026] The present invention allows a novel way for the isolation of DNA sequences from genetically related clones that are functionally different. Genetically related clones are those that belong to a particular structural class, for example immunoglobulins or alpha-beta-barrels. The invention allows for both isolation and random combination into a given DNA sequence of functional sequences from these related clones. These functional sequences may be loops that perform binding or catalysis.

[0027] The concept of the invention is demonstrated using antibody molecules where CDR-regions from different germline sequences can be isolated and randomly combined into a defined framework sequence. The invention expands the complexity of the molecular libraries that can be selected using phage display. The concept of the invention is also demonstrated by the affinity maturation of antibody fragments by the isolation and random combination of mutated CDR-regions.

[0028] It is not possible to use the DNA shuffling concept (Stemmer, 1994) to isolate specific sequences and randomly combine these into a given gene sequence, as it is not possible to amplify individual DNA regions formed in vivo using DNA shuffling. Combination of entire gene sequences is possible, but here defined regions cannot be shuffled. Rather all the DNA is shuffled. Thus, DNA sequences from genetically related clones that are functionally different, eg proteins that belong to structural classes like immunoglobulins or alpha-beta-barrels, cannot be shuffled in such a way that specific regions are kept constant and other regions are shuffled.

[0029] The system provided by the present invention offers a simple way to randomly combine functional regions of proteins (eg loops) to a defined (specifically selected) scaffold, ie shuffling of loops to a given protein tertiary structure in order to find new protein functions. Furthermore, the DNA shuffling technology introduces mutations at a rate of 0.7% (Stemmer, 1994). Thus, the known DNA shuffling technology (Stemmer, 1994) does not allow for shuffling of unmutated regions, since the process itself introduces mutations at random positions, including the scaffold regions.

[0030] In contrast, the invention allows for mutagenesis of defined DNA-sequences together with shuffling and assembly of these pieces of DNA into a coding region, and will allow for mutagenesis of defined regions and subsequent selection of these regions in combination.

[0031] The invention allows for different regions of DNA from different sequences (clones) to be shuffled and randomly combined. This increases the genetic variation from which functional antibody fragments are selected and will thus increase the probability of selecting proteins with the desired characteristics. It can be realised that by randomly shuffling as few as a hundred CDRs at each position in the VH and VL of an fragment, as many as 10^{12} combinations may be obtained thereby extending the variability normally found in the immune system.

[0032] The invention provides amplification of defined regions from eg a cDNA library using two primers of which one is biotinylated. Using the MSBP, e.g. biotin, group, single stranded DNA can be isolated and used in the gene assembly process. The present inventors have demonstrated this with the amplification of diverse CDR regions from an antibody gene library and the combination of these CDR regions randomly to a given framework region. Thus, defined regions of DNA (framework regions) can be interspaced by random regions of DNA (CDR regions), which have an in vivo origin or can be chemically synthesized.

[0033] The present invention also provides polynucleotide sequences and the proteins they encoded produced by the method described above. There is also provided vectors incorporating the polynucleotide sequences and host cell transformed by the vectors.

[0034] The present invention also provides a polynucleotide library comprising polynucleotides created by the method described above which may be used for phage display.

[0035] Aspects and embodiments of the present invention will now be illustrated, by way of example, with reference to the accompanying figures. Further aspects and embodiments will be apparent to those skilled in the art. All documents mentioned in this text are incorporated herein by reference.

Brief description of the drawings

[0036]

Figure 1 shows shuffling of specific DNA sequences between different clones, based on the assembly of gene sequences from a set of overlapping oligo-nucleotides following a one-step PCR protocol.

Figure 2 shows different dissociation rate constants for different CDR-shuffled clones. A low bar represents slow dissociation-rate, a high bar represents a fast dissociation-rate. Clone 36 is the original non-mutated antibody fragment.

Figure 3 shows the results of affinity purified scFv antibody fragment assayed on HPLC, Superose S-200 FPLC-column (Pharmacia) in PBS buffer. Peak 1 is the monomeric form of the antibody fragment, peak 2 is a small amount of impurity and peak 3 is NaN3 (sodium azid), used as a preservative.

Figure 4 shows a schematic representation of amplification of defined sequences of DNA and the shuffling of these into a master framework. Only the CDR regions are amplified. Figure 4A: Assembly of genes for the VH-domain. The template is scFv-B11 mutated with error prone PCR. An individual CDR is amplified using two primers adjacent to the particular CDR and one of these primers is biotinylated at the 5' end. The individual CDR is amplified and double-stranded DNA (dsDNA) is produced with the mutations focused to the CDR since the two amplification primers do not contain any mutations. This DNA is separated into two single stranded DNA molecules. The molecule without biotin is used in gene assembly. Primers 725, 729, 730, 728, 727 are synthesized in a DNA synthesizer and primers H2, H3, H5 contain mutated CDR and are amplified as above. Figure 4B: Assembly of genes for the VL-domain. CDRs are amplified in the same way as in A. Primers 759, 738, 745, 744, 880 are synthesized in a DNA synthesizer and primers L2, L3, L5 contain mutated CDR and are amplified as above.

Figure 5 shows the alignment of the peptide sequences for clones 3, 11 and 31 with the original non-mutated antibody fragment (wt). The CDR-regions are marked. Mutations in clones 3, 11 and 31 are underlined.

Figure 6 shows the principles for the isolation of single-stranded DNA for the shuffling of defined DNA regions.

Figure 7 shows the length of CDR3 heavy chain from different clones. These CDR regions have been amplified from different germline sequences and randomly cloned to a defined framework region (from DP-47 sequence).

Figure 8 shows a schematic representation of amplification of defined sequences of DNA and the shuffling of these into a master framework. All the oligonucleotides used in the gene assembly are amplified by PCR, but only the CDR regions contain any genetic variation. Figure 8A: Assembly of genes for the VH-domain. The template for the framework region amplification is scFv-B11, whereas CDRs are amplified from cDNA prepared from peripheral blood lymphocytes, tonsils and spleen. An individual DNA fragment is amplified using two primers located at the ends of the fragments to be amplified and one of these primers is biotinylated at the 5' end. The individual DNA fragment is amplified and double-stranded DNA (dsDNA) is produced. This DNA is separated into two single stranded DNA molecules. The molecule without biotin is used in gene assembly, i.e. primers H1, H4, H6 and these primers contain no variation. Primers HCDR1, HCDR2, HCDR3 contain different CDR and are amplified using two primers adjacent to the particular CDR and one of these primers is biotinylated at the 5' end. The individual CDR is amplified and double-stranded DNA (dsDNA) is produced with the variation focused to the CDR since the two amplification primers do not contain any mutations. This DNA is separated into two singled stranded DNA molecules and used in gene assembly of VH domain in a library format, i.e. the variation in the CDRs is derived from different germ-line sequences. Primers BT25 and BT26 are synthesized in a DNA-synthesizing machine. Figure 8B: Assembly of genes for the VL-domain. In principle the same procedure as in A. Primers L1, L4, L6 are amplified and produced by PCR and contain no variation. LCDR1, LCDR2, LCDR3 contain different CDR. Primers BT7 and BT12 are synthesized in a DNA-synthesizing machine.

Figure 9 shows the variation in a library constructed according to Fig. 8. The scFv region of library clones and original scFv-B11, binding to FITC (fluorescein-iso-thiocyanate) was synthesized by PCR. Purified PCR products were cut with BstNI and separated on a 2.5% agarose gel. Clones 1-15 are in lane 2-16, clones 16-29 are in lane 18-31. Original scFv-B11 is in lane 32. Analysis revealed that 28 clones could be sorted in 13 different groups according to restriction pattern and fragment size. Eight clones (1, 2, 8, 10, 12, 16, 26, 27) were unique, 2 clones (17, 24) appeared similar, 1 group of clones (18, 23, 29) had 3 similar members, 2 groups (5, 15, 14, 19) and (3, 4, 6, 11) had 4 members and 1 group (7, 9, 13, 20, 21, 22, 25) had 7 similar members. This experiment underestimates the variation in the library since BstNI detects only a fraction of sequence variability. In addition, the gel resolution did not allow the detection of minor size differences and did not resolve fragments below 100 bp.

Figure 9B shows clones showing similar restriction pattern in the experiment exemplified in Figure 9A cut by both BstNI and BamHI and separated on 3% agarose gels. To facilitate comparison, the groups of similar clones described in experiment A were put together on the gels. Clone 8 and 28 from experiment A were excluded due to space limitations.

Gel I) Lane 1-8; standard, clone 5,15,14,19,2,27, original scFv-B11, respectively
Gel II) Lane 1-8; standard, clone 16,17,24,18,23,29,26, respectively
Gel III) Lane 1-8; standard, clone 7,9,13,20,21,22,25, respectively
Gel VI) Lane 1-8; standard, clone 3,4,6,11,1,10,12, respectively

[0037] Under these improved experimental conditions, essentially all clones had different restriction patterns/fragments sizes. All clones were different from the original scFv-B11 gene (lane 8, gel 1). Moreover, the groups of clones which appeared similar in Figure 9A were found to be different as analyzed in Figure 9B. See clone 5,15,14,19 (lanes 2-5 gel I), clone 17,24 (lanes 3-4 gel II), clone 18,23,29 (lanes 5-7 gel II), clones 7,9,13,20,21,22,25, (lanes 2-8, gel III) and clones 3,4,6,11 (lanes 2-5 gel IV).

[0038] In conclusion, these experiments suggest that the library contains high variability.

Detailed description and exemplification of the invention

[0039] One aspect of the DNA shuffling procedure can be illustrated by the following steps in Fig 1.

A: A gene coding for a protein of interest is divided into overlapping oligonucleotides.
B: The oligonucleotides are assembled using PCR into a full length gene-sequence.
C: The gene sequence is subjected to mutagenesis, eg by error-prone PCR.
D: Pairs of oligonucleotides are synthesized, each pair covering a region defined by one of the oligonucleotides in step A above, except for a region located in the middle of the step A oligonucleotide. This uncovered region is the DNA sequence that can be shuffled after PCR amplification. These two synthesised oligonucleotides can thus be used as amplification primers to amplify the uncovered region.
E: One of these amplification primers is biotinylated and the double-stranded PCR product can then be isolated using well-known strepavidin systems.
F: From the thus isolated amplified oligonucleotides can be obtained a single-stranded DNA sequence containing DNA from the uncovered region mentioned above, which can then be used as oligo-nucleotide in a new assembly of the gene sequence as described in step A.
G: If DNA sequences from different clones and from different regions of the mutated gene sequence are amplified and made single-stranded, they will combine randomly in the PCR process of gene assembly. This random combination is the basis for in vitro molecular evolution.

Examples

[0040] The present inventors have demonstrated the concept of shuffling of defined DNA in different experimental settings. Firstly, the shuffling of in vitro mutated CDR regions in an antibody fragment for affinity maturation purposes (example 1 and 2) is exemplified and secondly the shuffling of in vivo formed CDRs for creation of a highly variable antibody library (example 3 and 4) is exemplified.

1. Affinity maturation

[0041] A model system was developed, based on the scFv-B11 antibody fragment which binds to FITC. The full-length gene encoding this scFv was assembled from a set of 12 oligonucleotides (Fig. 4A and Fig. 4B) representing the known DNA sequence of the scFv-B11, and the functional binding of the gene product to FITC could be verified. This gene sequence was then mutagenised using error-prone PCR, and the DNA encoding the CDR regions were amplified as described above, using the amplification primers, one of which is biotinylated. (The CDR regions are the parts of the antibody molecule involved in binding the antigen, in this case FITC).

[0042] All six CDR regions were amplified and a new gene was assembled using six oligonucleotides selected from the first assembly of 12 oligonucleotide (see above) (these were not mutagenized) and six from the amplification of mutagenized CDR regions. Selection of functional antibody fragments that bound FITC was carried out using phage display. 50% of the clones bound FITC with different dissociation-rates than did the original scFv-B11, as measured in the BIAcore biosensor (Figure 2). This demonstrates that the clones were changed in the way they recognized FITC.

[0043] Of the 16 clones identified to bind FITC in BIAcore (Figure 2) clones 3, 11, 27 and 31 were chosen to be analyzed in more detail as these clones exhibited the larger changes in off-rates. These clones were expressed and affinity-purified on a column conjugated with FITC-BSA and eluted with a low pH buffer. The purified scFv-antibody fragments were further purified and analyzed with HPLC, using a Pharmacia Superdex 200 FPLC column with the capacity to separate the monomeric and dimeric form of the antibodies. In all clones the monomeric form dominated (typical size profile is shown in Figure 3). This was then purified and used in detailed analysis of affinity using a BIAcore biosensor (Table 1).

Table 1.

Affinity determination of selected.			
Clone	$k_{ASS}(M^{-1}\ s^{-1})$	$k_{DISS}\ (s^{-1})$	$K_A\ (M^{-1})$
#3	$2,0 \times 10^5$	$4,3 \times 10^{-3}$	$4,8 \times 10^7$
#11	$2,6 \times 10^5$	$3,3 \times 10^{-3}$	$7,8 \times 10^7$
#27	$5,0 \times 10^5$	$16,0 \times 10^{-3}$	$3,1 \times 10^7$
#31	$1,2 \times 10^5$	$5,4 \times 10^{-3}$	$2,1 \times 10^7$
(FITC-B11 original)	$2,7 \times 10^5$	$9,7 \times 10^{-3}$	$2,8 \times 10^7$

[0044] Clone #11 exhibited an affinity 2.8 times higher than the original scFv-B11 antibody fragment. This increase is based on a slower off-rate. One clone (#27) showed 2 times increase in association-rate. However, the overall affinity of this clone was similar to the original FITC-B11 clone due to a faster dissociation-rate. The distribution of different association and dissociation-rates among the clones was considered a source for CDR-reshuffling for further improvement of affinities.

[0045] Three clones were sequenced. In the VH region (ie half of the scFv-B11 and carrying three CDR regions) the mutations found were all in the CDR regions as expected, since these were the only regions mutagenized and amplified using the amplification primers. Interestingly, all the CDR regions were different and carried different mutations (Figure 5). However, in the case of CDR region 2, the same mutation was found (a tyrosine to histidine substitution) in all 3 clones (the rest of CDR regions differed between the clones).

[0046] Furthermore, the mutation rates were found to be in between 2% and 4%, as determined from the base changes in the 90 bp long sequence built up from three CDR regions together. This is more than the error-prone PCR mutation rate, and indicates that there is combination of individual CDR regions from different clones.

2. Affinity maturation-reshuffling

[0047] In order to perform a second shuffling (reshuffling), clones selected for their binding affinity to FITC were used in an additional round of CDR-amplification and library construction. In theory, the reshuffled library will contain mutated shuffled CDR-regions, selected for improved binding to FITC. In this way, new combinations of CDR-regions, improved with respect to binding, could be constructed and the library subjected to selection for binders with improved affinities.

[0048] The pool of all clones obtained from the selection procedure (as detailed in example 1) were used as template for CDR amplifications. One amplification was carried out for each CDR using primers listed in Table 2.

Table 2

Sequences for primers used in CDR-shuffling.

B=Biotin labeled 5´ primer

CDR Reamplification Primers

764	5′ B-GTC CCT GAG ACT CTC CTG TGC AGC CTC TGG ATT CAC CTT T 3′
875	5′ TCC CTG GAG CCT GGC GGA CCC A 3′
876	5′ CGC CAG GCT CCA GGG AAG GGG CTG GAG TGG GTC TCA 3′
765	5′ B-GGA ATT GTC TCT GGA GAT GGT GAA 3′
799	5′ GAG CCG AGG ACA CGG CCG TGT ATT ACT GTG CAA GA 3′
766	5′ B-GCG CTG CTC ACG GTG ACC AGG GTA CCT TGG CCC CA 3′
767	5′ B-AGC GTC TGG GAC CCC CGG GCA GAG GGT CAC CAT CTC TTG T 3′
800	5′ GGG CCG TTC CTG GGA GCT GCT GGT ACC A 3′
801	5′ GCT CCC AGG AAC GGC CCC CAA ACT CCT CAT CTA T 3′
768	5′ B-GAC TTG GAG CCA GAG AAT CGG TCA GGG ACC CC 3′
802	5′ CTC CGG TCC GAG GAT GAG GCT GAT TAT TAC TGT 3′
769	5′ B-CGT CAG CTT GGT TCC TCC GCC GAA 3′

Framework VH

727	5′ CCG CCG GAT CCA CCT CCG CCT GAA CCG CCT CCA CCG CTG CTC ACG GTG ACC A 3′
728	5′GAC CGA TGG ACC TTT GGT ACC GGC GCT GCT CAC GGT GAC CA 3′
729	5′ GAG GTG CAG CTG TTG GAG TCT GGG GGA GGC TTG GTA CAG CCT GGG GGG TCC CTG AGA CTC TCC TGT 3′
730	5′ GGC CGT GTC CTC GGC TCT CAG GCT GTT CAT TTG CAG ATA CAG CGT GTT CTT GGA ATT GTC TCT GGA GAT GGT 3′

Framework VL

738	5′ CAG TCT GTG CTG ACT CAG CCA CCC TCA GCG TCT GGG ACC CCC G 3′
744	5′ ACT AGT TGG ACT AGC CAC AGT CCG TGG TTG ACC TAG GAC CGT CAG CTT GGT TCC TCC GC 3′
745	5′ CTC ATC CTC GGA CCG GAG CCC ACT GAT GGC CAG GGA GGC TGA GGT GCC AGA CTT GGA GCC AGA GAA TCG 3′
1129	5′ CAG GCG GAG GTG GAT CCG GCG GTG GCG GAT CGC AGT CTG TGC TGA CTC AGC CAC CCT CAG CGT CTG GGA CCC CCG 3′

Amplification primers VH/VL Assembly

1125	5′ ACT CGC GGC CCA ACC GGC CAT GGC CGA GGT GCA GCT GTT GGA G 3′
1126	5′ CAA CTT TCT TGT CGA CTT TAT CAT CAT CAT CTT TAT AAT CAC CTA GGA CCG TCA GCT TGG T 3′

[0049] The amplification was peformed according to following parameters: 100 ng template (1.6 x 10^8 CFU bacteria grown for 6 h), 60 pmol each primer, 5 Units PFU polymerase (Stratagene), 1 x PFU buffer, 500 μM dNTPs, reaction volume 100 μl, preheat 96°C for 10 minutes, 96°C for 1 minute: 68°C for 1 minute: 72°C for 1 minute for 25 cycles, 72°C for 10 minutes. This procedure was essentially the same as for CDR amplification in Example 1. The amplified CDR were used for assembly into VH and VL encoding sequence according to Figure 1, 4A, 4B and Table 3.

Table 3

PCR parameters for the assembly of VH and VL gene sequences in CDR-shuffling		
VL	VH	
Primer 759	Primer 725	30 pmol
Primer 738	Primer 729	0.6 pmol
Primer L2	Primer H2	0.6 pmol
Primer L3	Primer H3	0.6 pmol
Primer 745	Primer 730	0.6 pmol
Primer L5	Primer H5	0.6 pmol
Primer 744	Primer 728	0.6 pmol
Primer 880	Primer 727	30 pmol
Taq	Taq	10 Units
dNTPs	dNTPs	200 μM
1x Taq buffer	1x Taq buffer	to 100 μl
Preheat 95° 10 minutes, 20 cycles: 95° 1 minutes, 68° 1 minutes, 72°1 minutes 72°10 minutes.		

[0050] The VH and VL were then assembled into a scFv encoding sequence according to standard procedures (Griffiths et al 1994). The resulting library was subjected to panning so as to select binders with improved affinities to FITC. The selection procedure for the reshuffled library was essentially the same as for the initially shuffled library. The total number of clones obtained after selection was 510. Six clones (B) were chosen from this new pool and were tested and compared to 6 clones (A) from the first pool, originating from the shuffled library (Table 4).

Table 4

Dissociation-rates of individual clones selected from the shuffled library (clones A) and from the reshuffled library (clones B).	
Clone	K_{DISS} (s-1 x10^{-3})
scFv-B11 (original)	12.9
1A	6.3
12A	5.7
13A	9.0
14A	9.7
16A	1.8
17A	7.9
22B	0.2
31B	0.3
32B	9.8
33B	6.8
34B	7.3
35B	8.7

[0051] Two clones from the reshuffling experiments (22B and 31B) exhibited substantially slower dissociation-rates, indicating that the reshuffling process yielded binders with improved affinities.

3.Cloning and shuffling of defined DNA regions

[0052] In our system it is possible to amplify defined regions from a cDNA library using two primers of which one is biotinylated. Using the biotin group, single stranded DNA can be isolated an used in the gene assembly process (Figure 6). We have demonstrated this with the amplification of diverse CDR regions from an antibody gene library and the combination of these CDR regions randomly to a given framework region. Thus, defined regions of DNA (framework regions) can be interspaced by random regions of DNA (CDR regions) which have an in vivo origin (Table 5). The CDR3 region vary in size (Figure 7). Alternatively, these regions could be chemically synthesised.

Table 5

Combination of CDR regions from different germline sequences transplanted to the DP-47 framework encoding the variable heavy domain. For CDR1 and CDR2 the suggested germline origin is indicate. For CDR3 the number of residues in the CDR-region is written. N.D = not determined.

Clone	CDR1	CDR2	CDR3
1	DP-35	DP-42	12
2	DP-49	DP-53	13
3	N.D.	DP-51	11
4	DP-32	DP-47	10
5	DP-41	DP-47	8
6	DP-32	DP-77	9
7	DP-31	DP-47	7
8	DP-49	DP-35	5
9	DP-49	DP-35	N.D.
10	DP-48	DP-48	N.D.
11	DP-51	DP-47	10
12	DP-34	DP-31	N.D.
13	DP-85	DP-53	4
14	DP-31	DP-77	10
15	DP-34	DP-53	4

4. Library construction.

[0053] A gene library was constructed encoding scFv antibody fragments. The strategy used for this library is based on the assembly of a set of oligonucleotides into a sequence encoding VH and VL antibody domains (Figure 8A, 8B.) Native *in vivo* formed CDR regions can be shuffled and assembled into a given master framework. In this example we have developed this concept further and assembled both VH and VL encoding gene sequences with native CDR regions into a given master framework. Thus, all six CDR positions have been shuffled. The template origin for CDR amplification was cDNA from peripheral blood B-cells, spleen, tonsills and lymphnodes. Oligonucleotides encoding the framework regions have also been amplified using the strategy with two flanking primers, where one is biotinylated (primers L1,H1 L4, H4, L6, H6). The primers used are described in Table 6 and in Figure 8A, 8B.

Table 6

Sequences for primers used in library construction.

B = Biotin labeled 5´ primer

Amplification of framework fragments

BT1. 5′ ACA GTC ATA ATG AAA TAC CTA TTG C 3′
BT2. 5′ B-GC ACA GGA GAG TCT CA 3′
BT3. 5′ B-CA CCA TCT CCA GAG ACA ATT CC 3′
BT4. 5′ GGC CGT GTC CTC GGC TCT 3′
BT5. 5′ B-TG GTC ACC GTG AGC AGC 3′
BT6. 5′ CCG CCG GAT CCA CCT 3′
BT7. 5′ CAG GCG GAG GTG GAT CCG GC 3′

BT8. 5' B-CG GGG GTC CCA GAC GCT 3'
BT9. 5' B-CG ATT CTC TGG CTC CAA GT 3'
BT10. 5' CTC ATC CTC GGA CCG GA 3'
BT11. 5' B-TC GGC GGA GGA ACC AAG CT 3'
BT12 5' TGG CCT TGA TAT TCA CAA ACG AAT 3'

Amplification of in vivo CDR

BT13. 5' B-TC CCT GAG ACT CTC CTG TGC AGC CTC TGG ATT CAC CTT 3'
BT14. 5' TTC CCT GGA GCC TGG CGG ACC CA 3'
BT15. 5' B-GG AAT TGT CTC TGG AGA TGG TGA A 3'
BT16. 5' GTC CGC CAG GCT CCA 3'
BT17. 5' B-CG CTG CTC ACG GTG ACC AGT GTA CCT TGG CCC CA 3'
BT18. 5' AGA GCC GAG GAC ACG GCC GTG TAT TAC TGT 3'
BT19. 5' B-AG CGT CTG GGA CCC CCG GGC AGA GGG TCA CCA TCT CTT 3'
BT20. 5' GGG CCG TTC CTG GGA GCT GCT GAT ACC A 3'
BT21. 5' GCT CCC AGG AAC GGC CCC CAA ACT CCT CAT CTA T 3'
BT22. 5' B-GA CTT GGA GCC AGA GAA TCG GTC AGG GAC CCC 3'
BT23. 5' B-GT CAG CTT GGT TCC TCC GCC GAA 3'
BT24. 5' CTC CGG TCC GAG GAT GAG GCT GAT TAT TAC T 3'

Assembly of VH and VL

BT25. 5' B-TA CCT ATT GCC TAC GGC AGC CGC TGG ATT GTT ATT ACT CGC GGC CCA GCC GGC CAT GGC CGA 3'
BT26. 5' CCG CCG GAT CCA CCT CCG CCT GAA CCG CCT CCA CCG CTG CTC ACG GTG ACC A 3'

Amplification primers 2nd assembly

BT27. 5' B-TGG CCT TGA TAT TCA CAA ACG AAT 3'
BT28. 5' B-ACG GCA GCC GCT GGA TTG 3'

[0054] The PCR parameters for CDR and framework region amplification were essentially the same as described in example 2. The PCR parameters for assembly of genes encoding VH and VL are described in Table 7.

Table 7

PCR parameters for the assembly of VH and VL gene sequences for library construction.		
VH	VL	
Primer BT25	Primer BT7	30 pmol
Primer H1	Primer L1	0.6 pmol
Primer HCDR1	Primer LCDR1	0.6 pmol
Primer HCDR2	Primer LCDR2	0.6 pmol
Primer H4	Primer L4	0.6 pmol
Primer HCDR5	Primer LCDR3	0.6 pmol
Primer H6	Primer L6	0.6 pmol

Table 7 (continued)

PCR parameters for the assembly of VH and VL gene sequences for library construction.		
VH	VL	
Primer BT26 Taq dNTPs 1x Taq buffer	Primer BT12 Taq dNTPs 1x Taq buffer	30 pmol 10 Units 200 μM to 100 μl
Preheat 95° 10 minutes, 20 cycles: 95° 1 minutes, 68° 1 minutes, 72° 1 minutes and 72° 10 minutes.		

[0055] The assembled VH and VL gene sequences were assembled into a scFv coding sequence using standard protocols (Griffiths et al 1994). A library of 1.1 x 10^9 members were constructed out of the 40 clones tested all 40 contained an insert of the right size as determined by PCR agarose gel electrophoresis. In order to test the variability in the library, PCR amplified and purified inserts were subjected to cleavage by BsTNI and BamHI. Clones showed different restriction patterns, as determined by agarose gel electrophoresis and compared to the control scFv-B11 (Figure 9).

[0056] In order to estimate the frequency of clones able to express scFv antibody fragments, clones from the library containing the FLAG sequence (Hopp et al, 1989), as well as control bacteria with and without FLAG sequence, were plated at low density on Luria broth-plates containing 100μg/ml ampicillin, 25μg/ml tetracycline and 1% glucose. The plates were grown at 37°C over night and lifted to nitrocellulose filters by standard methods (Sambrook et al 1989). In order to induce synthesis of the scFv genes in the bacteria, filters were incubated for 4hrs on plates containing 0.5mM isopropyl-thio-β-D-galactoside (IPTG) but without glucose. Bacteria were then lysed by lyzosyme/chloroform treatment, the filters were washed and incubated with anti-FLAG M2 antibody (Kodak) followed by anti-mouse peroxidase conjugated second antibody (P260 Dakopatts) and detected by DAB 3,3'-diaminobenzidine tetrahydroklorid, Sigma) (Table 8).

Table 8

Frequency of intact antibody genes in the library			
Library Pool	Tested clones	FLAG positive clones	Percent positive clones
A	145	88	60
B	77	52	67
C	158	105	66
D	68	48	70
All library pools	448	293	65.4
Positive control pFAB5cHis scFvB11	64	64	100
Negative control pFAB5cHis	30	0	0

[0057] The anti-FLAG antibody detects a FLAG sequence situated downstream of the scFv gene in the library constructs as well as in the control vector pFAB5cHis scFvB11, but not in the original vector pFAB5cHis. Clones, to which the anti-FLAG antibody binds, therefore contains an intact open reading frame of the scFv gene.

[0058] The present application, therefore, provides a method of obtaining a polynucleotide sequence encoding a protein of desired characteristics by incorporating variant peptide regions (variant motifs) into defined peptide regions (scaffold sequence), comprising the steps of

a) subjecting a parent polynucleotide sequence encoding one or more protein motifs to mutagenesis to create a plurality of differently mutated derivatives thereof, or obtaining a parent polynucleotide encoding one or more variant protein motifs;
b) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif, and using each said pair of oligonucleotides as amplification primers to amplify the intervening motif;
c) obtaining single-stranded nucleotide sequences from the thus-isolated amplified nucleotide sequences; and
d) assembling polynucleotide sequences encoding a protein by incorporating nucleotide sequences derived from

step c) above with nucleotide sequence encoding scaffold sequences.

[0059] The method may further comprise the step of expressing the resulting protein encoding by the assembled polynucleotide sequence and screening for desired properties. The oligonucleotides may be single stranded.
[0060] One of said pair of oligonucleotides may be linked to a member of a specific binding pair (MSBP). Accordingly, the method may further comprise the steps of isolating the amplified variant motif by binding the MSBP to its specific binding partner. The MSBP may be biotin, in which case the specific binding partner may be streptavidin.
[0061] The method as described above may include the step of subjecting the parent polynucleotide sequence to mutagenesis by error-prone PCR.
[0062] The parent polynucleotide sequence may encode an antibody or part thereof.
[0063] The present application also provides a polynucleotide sequence encoding a protein of desired characteristics obtained by a method as described above, such as an antibody or fragment thereof.
[0064] The present application also provides a vector comprising a polynucleotide sequence as described above.
[0065] The present application also provides a host cell transformed with the vector described above.
[0066] The present application also provides a method of producing a polypeptide of desired characteristics comprising culturing the host cell as described above so that the polypeptide is produced, and optionally recovering the polypeptide produced.
[0067] The present application also provides polynucleotide library comprising polynucleotide sequences as described above.
[0068] The present application also provides a protein having desired characteristics obtained by a method as described above.
[0069] The present application also provides a method of creating a polynucleotide library comprising the steps of

a) subjecting a parent polynucleotide sequence encoding one or more protein motifs to mutagenesis to create a plurality of differently mutated derivatives thereof, or obtaining a parent polynucleotide encoding one or more variant protein motifs;
b) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif, and using each said pair of oligonucleotides as amplification primers to amplify the intervening motif;
c) obtaining single-stranded nucleotide sequences from the thus-isolated amplified nucleotide sequences;
d) assembling polynucleotide sequences by incorporating nucleotide sequences derived from step c) above with nucleotide sequence encoding scaffold sequence; and
e) inserting said polynucleotide sequences into suitable vectors; and optionally
f) further comprising the step of screening the library for a protein of desired characteristics.

References

[0070]

Barbas, C F et al: Proc Natl Acad Sci USA, 88:7978-82 (1991)
Barbas, C F et al: Proc Natl Acad Sci USA, 89:4457-61 (1992)
Crameri, A et al: Biotechniques, 18:194-196 (1995)
Griffiths, A D et al: EMBO J, 13:3245-3260 (1994)
Hoogenboom, H R et al: J Mol Biol, 227:381-8 (1992)
Hopp, T.P. et al: Biotechniques 7: 580-589 (1989)
Huse, W D et al: Science, 246:1275-81 (1989)
Larrick, J W et al: Biochem Biophys Res Commun, 160:1250-6 (1989)
Leung, D W et al: Technique, 1:11-15 (1989)
Marks, J D et al: Biotechnology, 10:779-83 (1992)
McCafferty, J et al: Nature, 348:552-4 (1990)
Moore, J C et al: Nature Biotechnology, 14:458-467 (1996)
Parmley, S F et al: Gene, 73:305-318 (1988)
Roberts, S et al: Nature, 328:731-4 (1987)
Sambrook, J et al:Molecular cloning. A laboratory manual. Cold spring Harbor Laboratory Press 1989.
Stemmer, W P: Nature, 370:389-391 (1994)
Yang, W P et al: J Mol Biol, 254:392-403 (1995)

Claims

1. A method of obtaining a polynucleotide sequence encoding a protein of desired characteristics comprising the steps of incorporating at least one variant nucleotide region (variant motif) into defined nucleotide regions (scaffold sequence) derived from a parent polynucleotide sequence.

2. A method according to Claim 1 wherein the sequence of the variant nucleotide motif is unknown at the time of incorporation.

3. A method according to any preceding claim wherein the at least one variant nucleotide regions encode specified regions of a protein.

4. A method according to any preceding claim wherein the parent polynucleotide sequence is subjected to error-prone PCR.

5. A method according to any preceding claim comprising the steps of

   a) subjecting the parent polynucleotide sequence encoding one or more protein motifs to mutagenesis to create a plurality of differently mutated derivatives thereof;
   b) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif, and using each said pair of oligonucleotides for copying and amplification of the intervening variant motif; and
   c) assembling polynucleotide sequences encoding a protein by incorporating the variant motif nucleotide sequences derived from step b) above with nucleotide sequence encoding scaffold sequences.

6. A method according to any one of Claims 1 to 3 comprising the steps of

   a) obtaining a parent polynucleotide encoding one or more variant protein motifs;
   b) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif, and using each said pair of oligonucleotides for copying and amplification of the intervening variant motif; and
   c) assembling polynucleotide sequences encoding a protein by incorporating the variant motif nucleotide sequences derived from step b) above with nucleotide sequence encoding scaffold sequences.

7. A method according to Claim 5 or 6 further comprising the step of obtaining single stranded nucleotide sequences from the amplified variant motif produced in step b), which single stranded nucleotide sequences are used in assembling the polynucleotde sequences encoding a protein in step c).

8. A method according to any one of Claims 5-7 wherein each pair of oligonucleotides are complementary to regions of scaffold sequence adjacent to the intervening variant motif.

9. A method according to any one of Claims 5-8 wherein the oligonucleotides are single stranded.

10. A method according to any one of Claims 5-9 wherein one of said pair of oligonucleotides is linked to a member of a specific binding pair (MSBP).

11. A method according to Claim 10 further comprising the steps of isolating the amplified variant motif by binding the MSBP to its specific binding partner.

12. A method according to Claim 10 or 11 wherein the MSBP is biotin.

13. A method according to Claim 12 wherein the specific binding partner is streptavidin.

14. A method according to any one of the preceding claims wherein the polynucleotide sequences encoding a protein are assembled using PCR.

15. A method according to any one of the preceding claims wherein the assembled polynucleotide sequence encodes an antibody of part thereof.

16. A method according to Claim 15 wherein the assembled polynucleotide sequence encodes an antibody part and wherein the part is Fab, Fd, Fv, dAb, $F(ab')_2$ or an scFv fragment, or an isolated CDR.

17. A method according to Claim 15 or 16 wherein complementarity determining regions (CDRs) of the antibody or part thereof are encoded in the thus obtained polynucleotide by variant motifs.

18. A method according to any one of the preceding claims further comprising the step of expressing the resulting protein encoded by the assembled polynucleotide sequence.

19. A method according to Claim 18 further comprising the step of screening the thus expressed protein for desired properties.

20. A method according to Claim 18 or 19 wherein the protein is an antibody or part thereof.

21. A method according to Claim 20 wherein the protein is an antibody part and wherein the part is Fab, Fd, Fv, dAb, $F(ab')_2$ or an scFv fragment, or an isolated CDR.

22. A method of creating a polynucleotide library comprising the steps of mutating a specified region (variant motif) of a parent polynucleotide sequence encoding one or more protein motifs by mutagenesis to create a plurality of differently mutated variant motifs and incorporating these mutated variant motifs into defined nucleotide regions (scaffold sequence) of the original parent polynucleotide sequence.

23. A method according to Claim 22 wherein the sequence of the variant nucleotide motifs are unknown at the time of incorporation.

24. A method according to Claim 22 or 23 wherein the parent polynucleotide sequence is subjected to error-prone PCR.

25. A method according to Claim 22, 23 or 24 comprising the steps of

a) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequence bounding an intervening variant protein motif, and using each said pair of oligonucleotides for copying and amplification of the intervening motif;
b) assembling polynucleotide sequences by incorporating the variant motif nucleotide sequences derived from step a) above with nucleotide sequences encoding scaffold sequence; and
c) inserting said polynucleotide sequences into suitable vectors.

26. A method of creating a polynucleotide library comprising the steps of obtaining parent polynucleotides encoding one or more variant protein motifs and incorporating the polynucleotide regions encoding the one or more variant protein motifs into defined nucleotide regions (scaffold sequence).

27. A method according to Claim 26 wherein the sequences of the one or more variant nucleotide motifs are unknown at the time of incorporation.

28. A method according to Claim 26 or 27 wherein the one or more variant protein motifs are at specified regions of the protein.

29. A method according to any one of Claims 26-28 comprising the steps of

a) providing a plurality of pairs of oligonucleotides, each pair representing spaced apart locations on the parent polynucleotide sequences bounding intervening variant protein motifs and using each said pair of oligonucleotides for copying and amplification of the intervening variant motifs;
b) assembling polynucleotide sequences by incorporating the variant motif nucleotide sequences derived from step a) above with nucleotide sequences encoding scaffold sequence; and
c) inserting said polynucleotide sequences into suitable vectors.

30. A method according to Claim 25 or 29 further comprising obtaining single stranded nucleotide sequences from the amplified variant motif produced in step a), which single stranded nucleotide sequences are used in assembling the polynucleotide sequences encoding a protein in step b).

31. A method according to any one of Claims 25, 29 or 30 wherein each pair of oligonucleotides are complementary to regions of scaffold sequence adjacent to the intervening variant motif.

32. A method according to any one of Claims 25 or 29-31 wherein one of said pair of oligonucleotides is linked to a member of a specific binding pair (MSBP).

33. A method according to Claim 32 further comprising the steps of isolating the amplified variant motif by binding the MSBP to its specific binding partner.

34. A method according to Claim 32 or 33 wherein the MSBP is biotin.

35. A method according to Claim 34 wherein the specific binding partner is streptavidin.

36. A method according to any one of Claims 22-35 wherein the polynucleotide sequences encoding a protein are assembled using PCR.

37. A method according to any one of Claims 22-36 further comprising the step of expressing said polynucleotide sequences to obtain corresponding library of polypeptide sequences.

38. A method according to Claim 37 further comprising screening the library for a polypeptide of desired characteristics.

39. A method of creating a polynucleotide library comprising obtaining a polynucleotide that encodes a polypeptide selected by the method of Claim 38 and reshuffling one or more times said polynucleotide by a method as defined in any one of Claims 22-36.

40. A method according to Claim 39 further comprising the step of expressing the thus obtained polynucleotide library sequences to obtain corresponding library of polypeptide sequences.

41. A method according to Claim 40 further comprising screening the library for a polypeptide of desired characteristics.

42. A method according to any one of Claims 37, 38, 40 or 41 wherein the thus obtained polypeptide is an antibody or part thereof.

43. A method according to Claim 42 wherein the thus obtained polypeptide is an antibody part and wherein the part is Fab, Fd, Fv, dAb, $F(ab')_2$ or an scFv fragment, or an isolated CDR.

44. A polynucleotide obtainable by a method according to any one of Claims 1 to 17 or 39.

45. A vector comprising a polynucleotide according to Claim 38.

46. A cell comprising a polynucleotide according to Claim 44 or a vector according to Claim 45.

47. A protein encoded by a polynucleotide according to Claim 44 or obtainable by a method according to any one of Claims 18-21, 37, 38 or 35-43 having the desired characteristics.

48. A library of polynucleotides or polypeptides obtainable by a method according to any one of Claims 22-43.

49. A method according to any one of Claims 1-43 further comprising the step of sequencing the thus obtained polynucleotide.

Fig.1.

Shuffling of defined regions of DNA

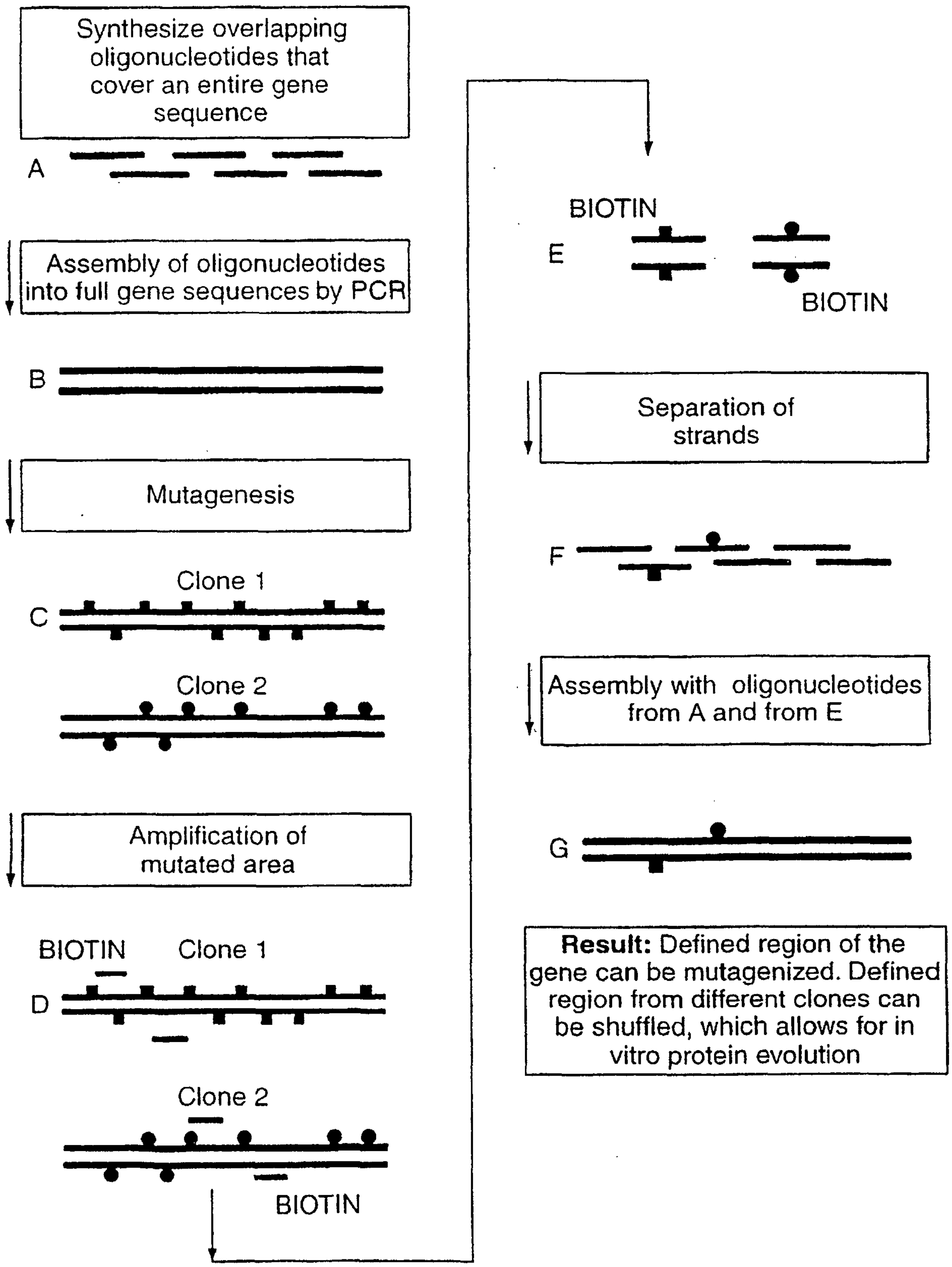


Fig.2.

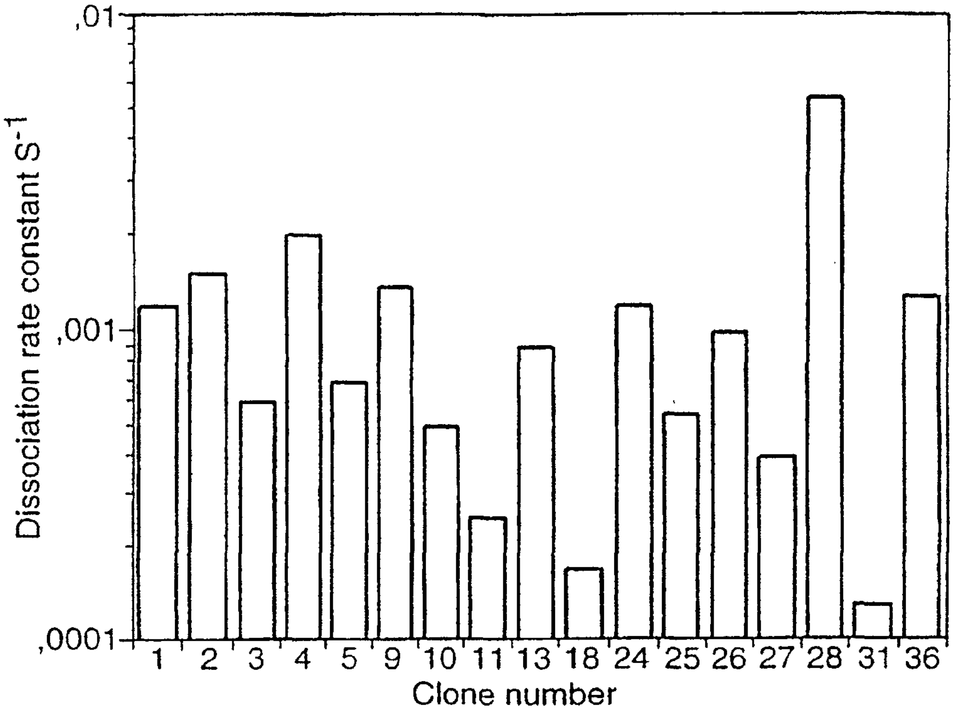

,01
,001
,0001
Dissociation rate constant S^{-1}
1 2 3 4 5 9 10 11 13 18 24 25 26 27 28 31 36
Clone number

Fig.3.

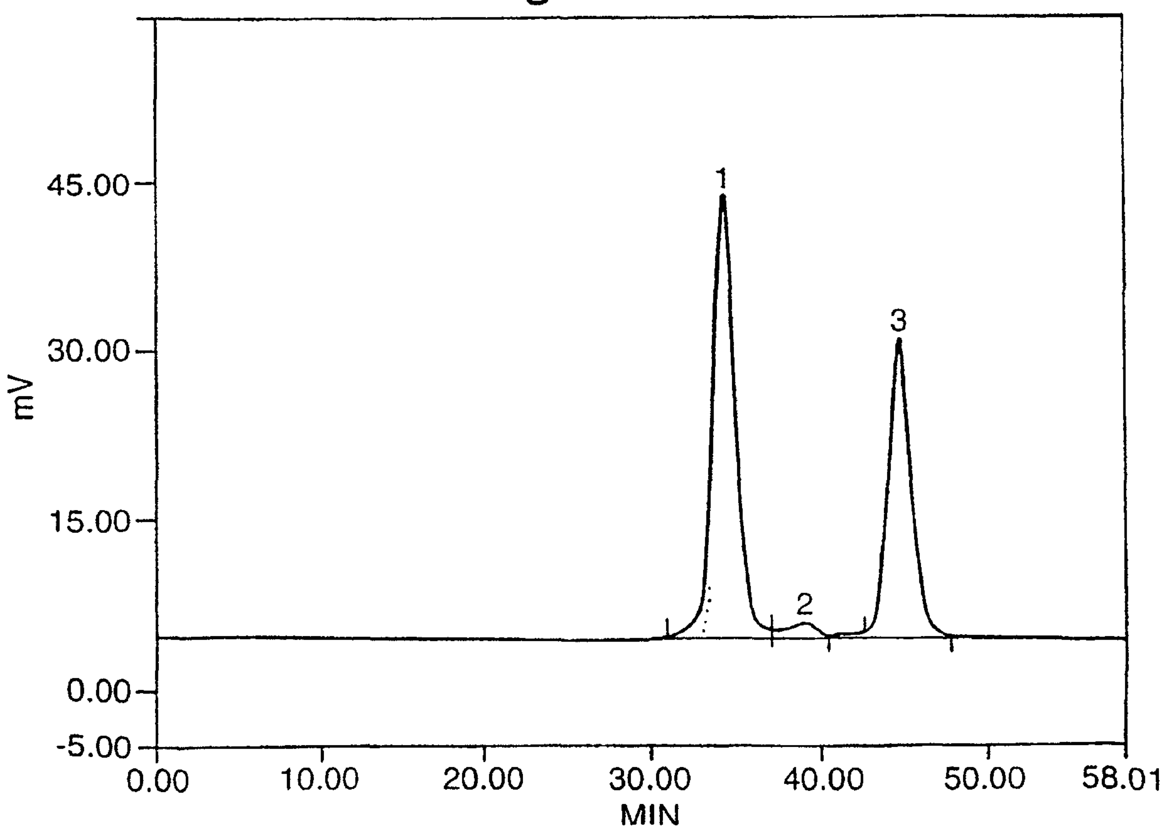
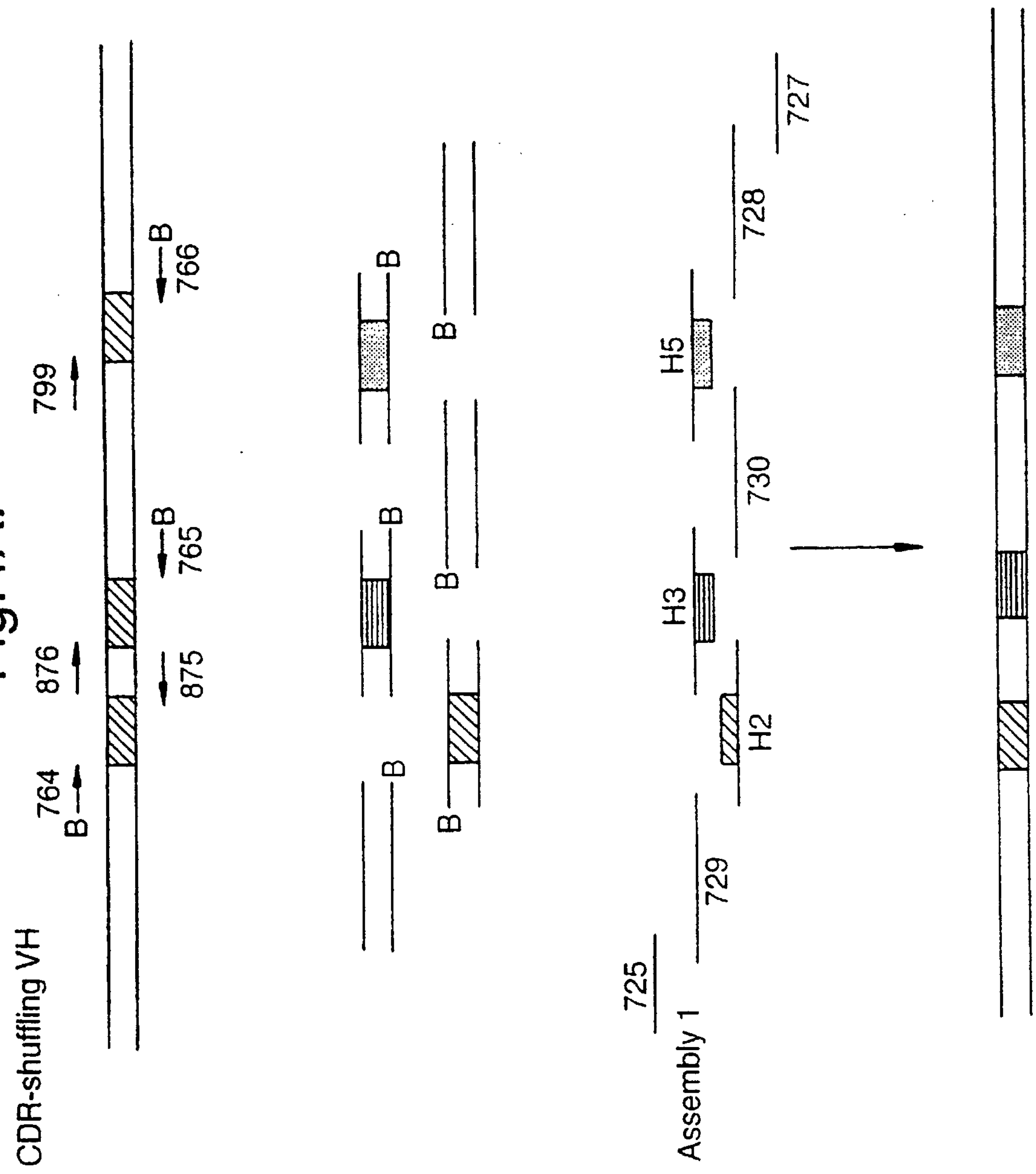

45.00
30.00
15.00
0.00
-5.00
mV
1
2
3
0.00 10.00 20.00 30.00 40.00 50.00 58.01
MIN

Fig.4A.


CDR-shuffling VH
764
B
876
875
765
B
799
766
B
Assembly 1
725
729
H2
H3
730
H5
728
727

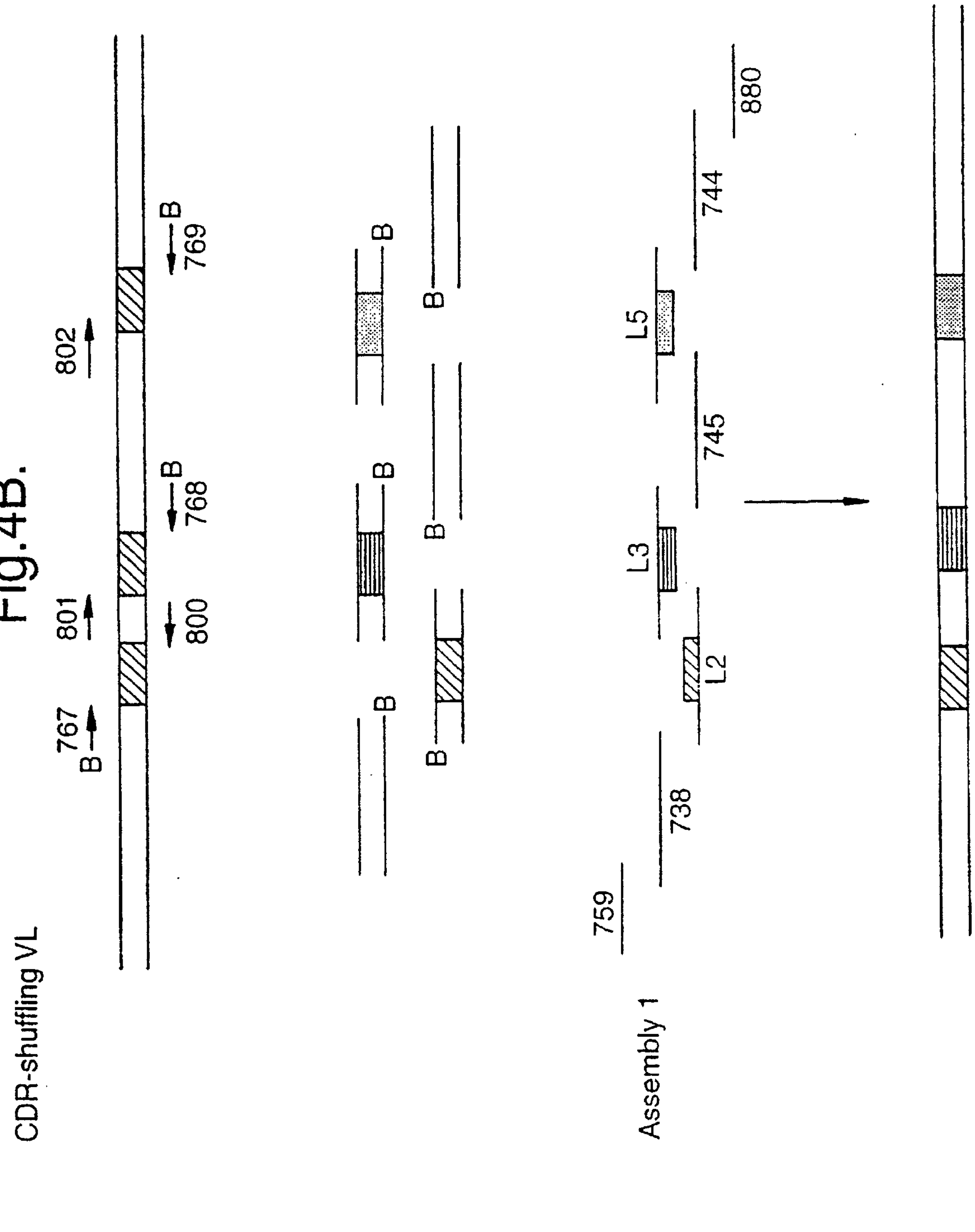

CDR-shuffling VL
767
B
801
800
768
B
802
769
B
B
B
B
B
B
B
Assembly 1
759
738
L2
L3
745
L5
744
880

Fig.4B.

Fig.5.

```
                                                       CDR1
CLONE 31   LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTFSRYAMSWVRQAPG
CLONE 3    LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTFSGYAMSWVRQAPG
CLONE 11   LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPG
ORIGINAL   LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPG

                     CDR2
CLONE 31   KGLEWVSAISGSGGGTHYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTA
CLONE 3    KGLEWVSAISGSGGSTHYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTA
CLONE 11   KGLEWVSAISGSGGSTHYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTA
ORIGINAL   KGLEWVSAISGSGGSTYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTA

                  CDR3
CLONE 31   VYYCARIGQFWGQGTLVTVSSGGGSGGGSQ
CLONE 3    VYYCARIGRFWGQGTLVTVSSGGGSGGGSQ
CLONE 11   VYYCARIGQFWGQGTLVTVSSGGGSGGGSQ
ORIGINAL   VYYCARIGQFWGQGTLVTVSSGGGSGGGSQ
```

Fig.6.

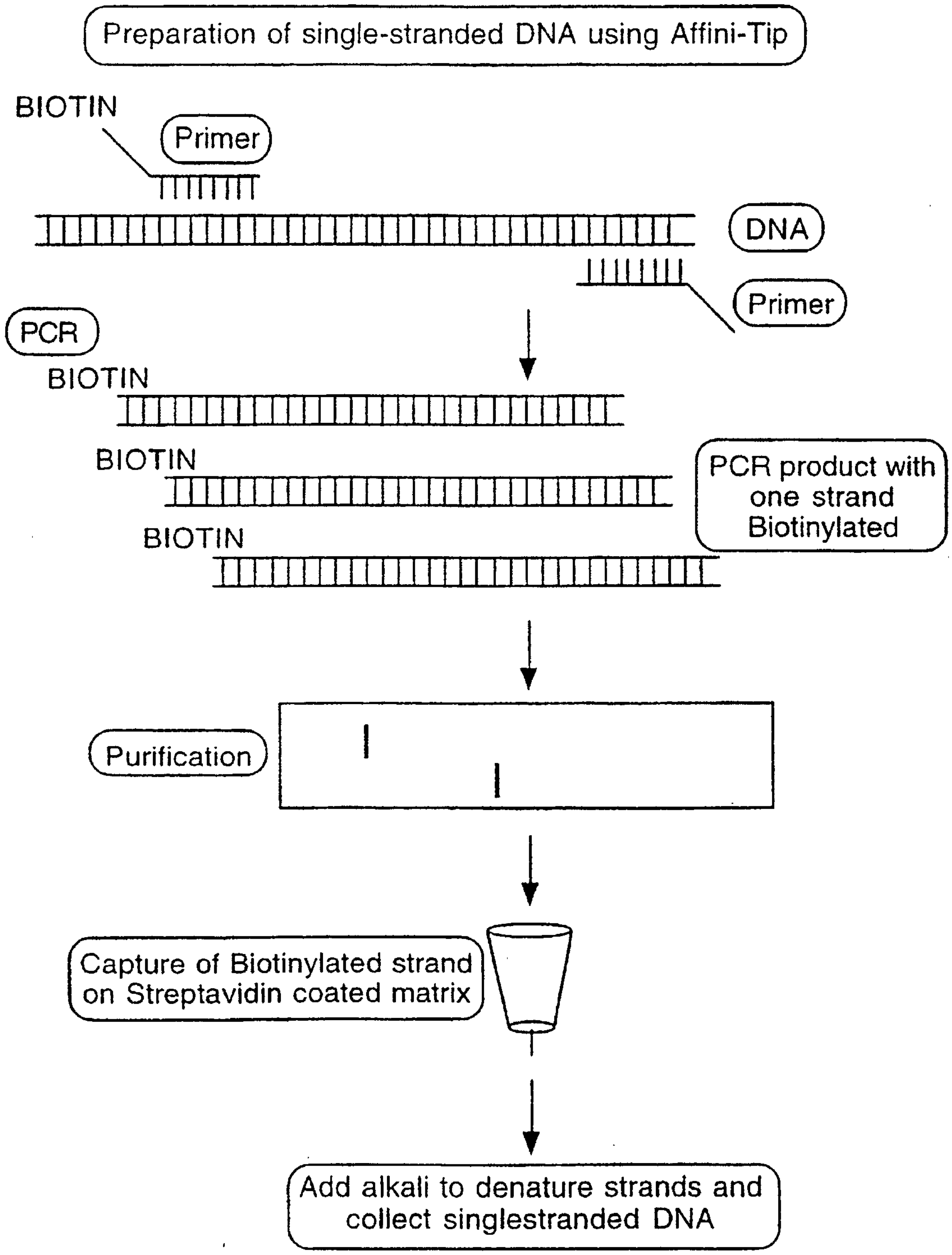

Preparation of single-stranded DNA using Affini-Tip
BIOTIN
Primer
DNA
Primer
PCR
BIOTIN
BIOTIN
BIOTIN
PCR product with one strand Biotinylated
Purification
Capture of Biotinylated strand on Streptavidin coated matrix
Add alkali to denature strands and collect singlestranded DNA

Fig.7.

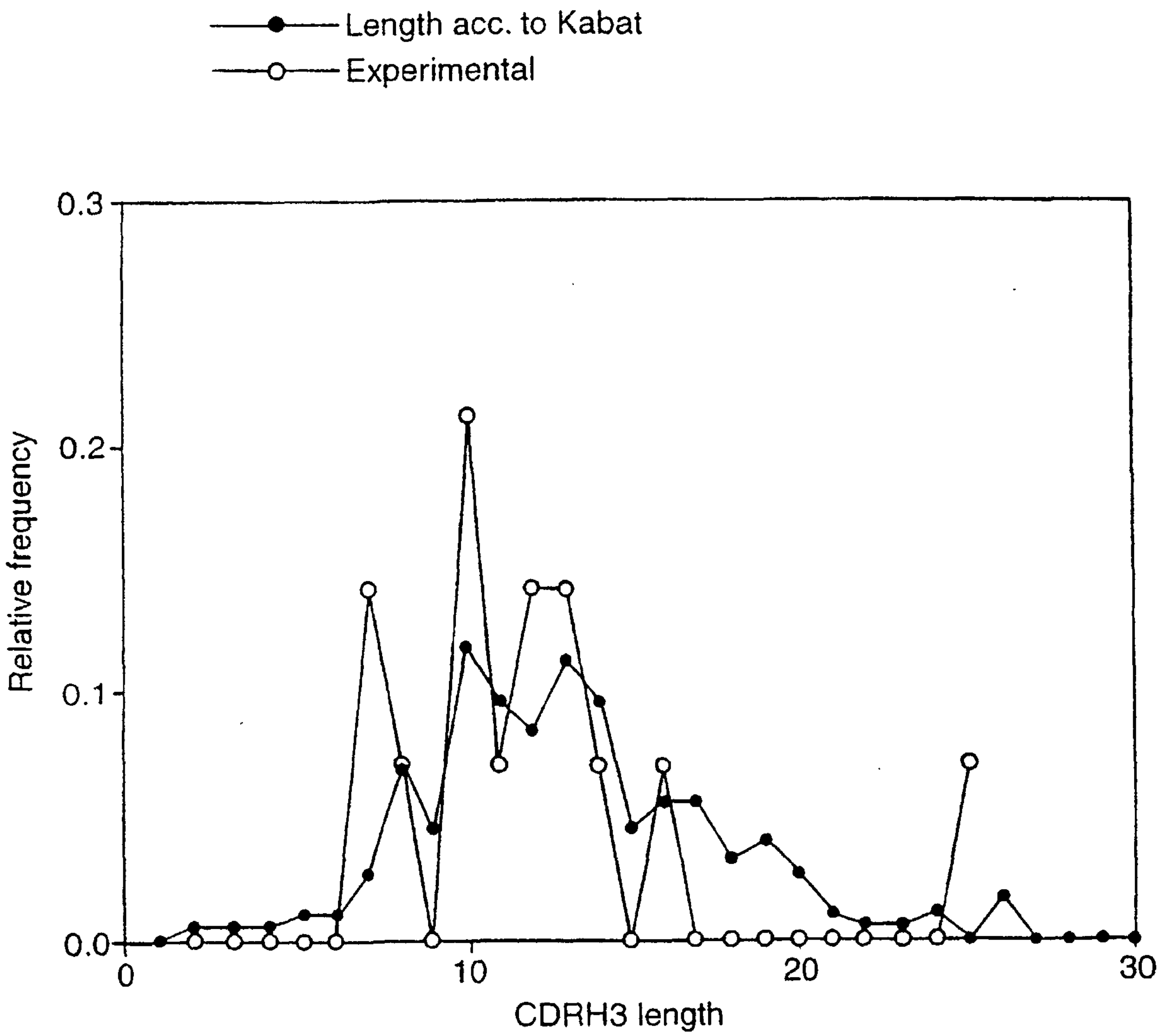

Length acc. to Kabat
Experimental
0.3
0.2
0.1
0.0
Relative frequency
0
10
20
30
CDRH3 length

Fig.8A.

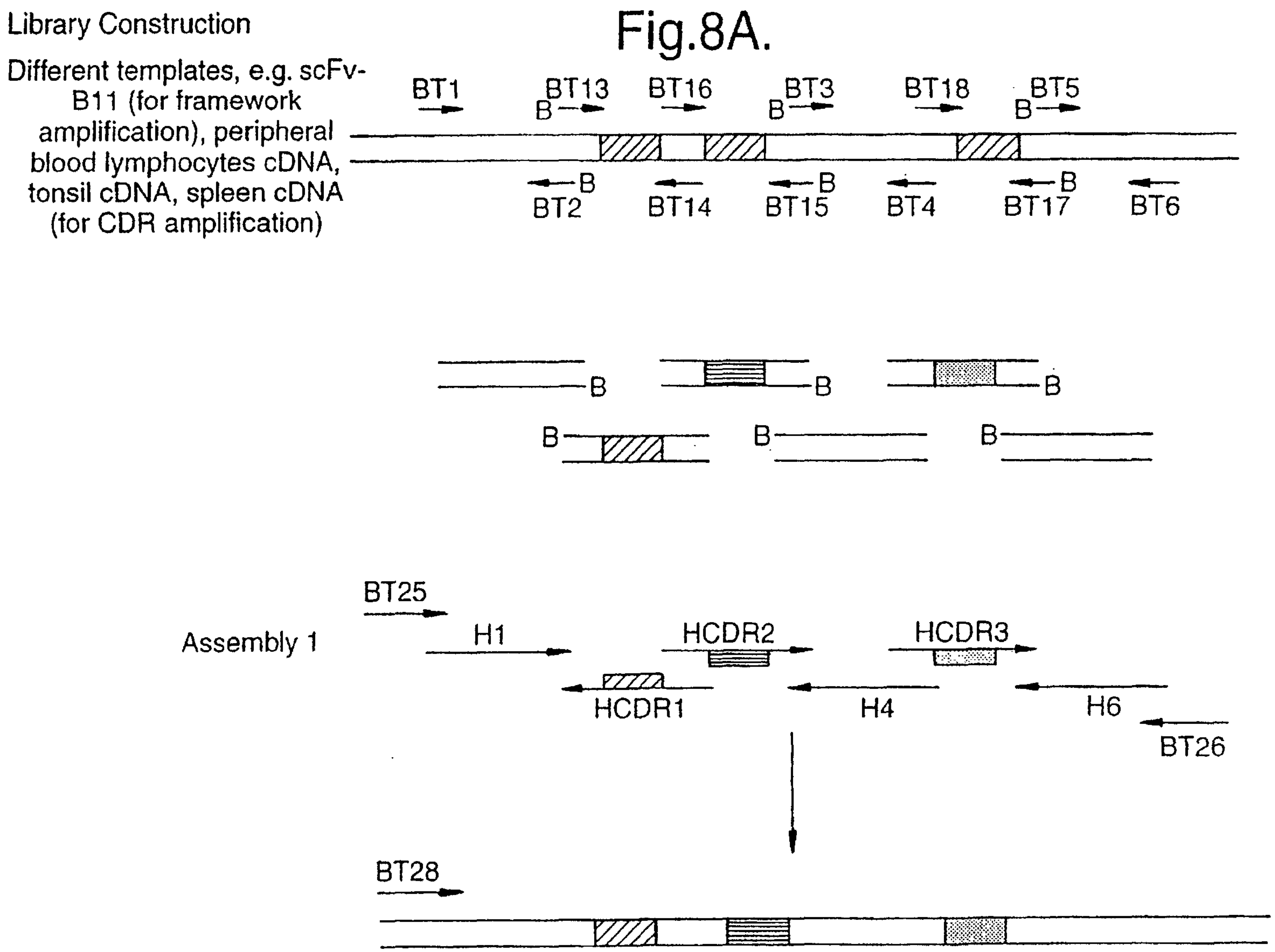

Library Construction
Different templates, e.g. scFv-B11 (for framework amplification), peripheral blood lymphocytes cDNA, tonsil cDNA, spleen cDNA (for CDR amplification)
BT1
BT13
BT16
BT3
BT18
BT5
B
BT2
BT14
BT15
BT4
BT17
BT6
Assembly 1
BT25
H1
HCDR2
HCDR3
HCDR1
H4
H6
BT26
BT28

Fig.8B.

Library Construction

Different templates, e.g. scFv-B11 (for framework amplification), peripheral blood lymphocytes cDNA, tonsil cDNA, spleen cDNA (for CDR amplification)

Assembly 1

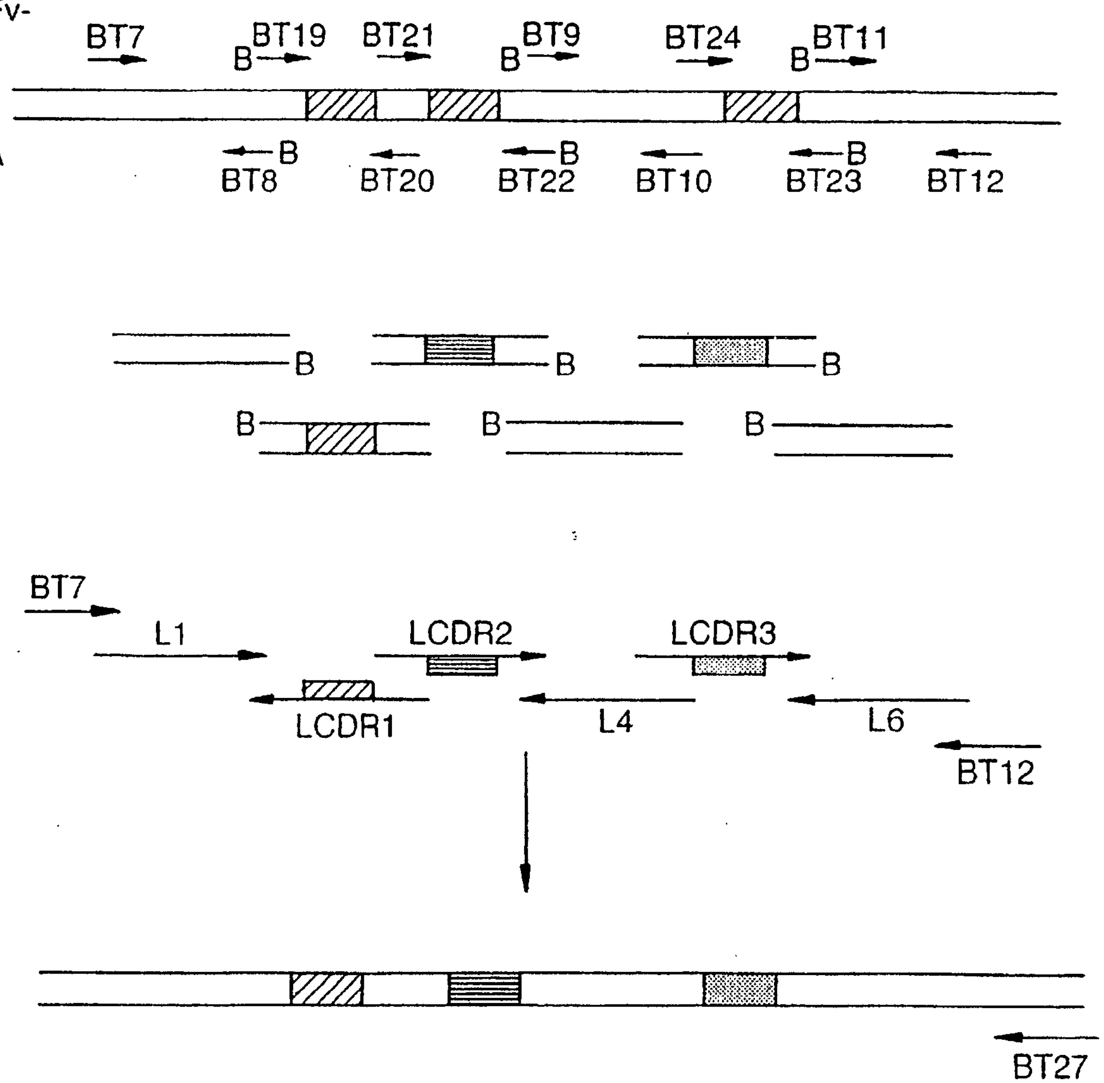


Fig.9A.

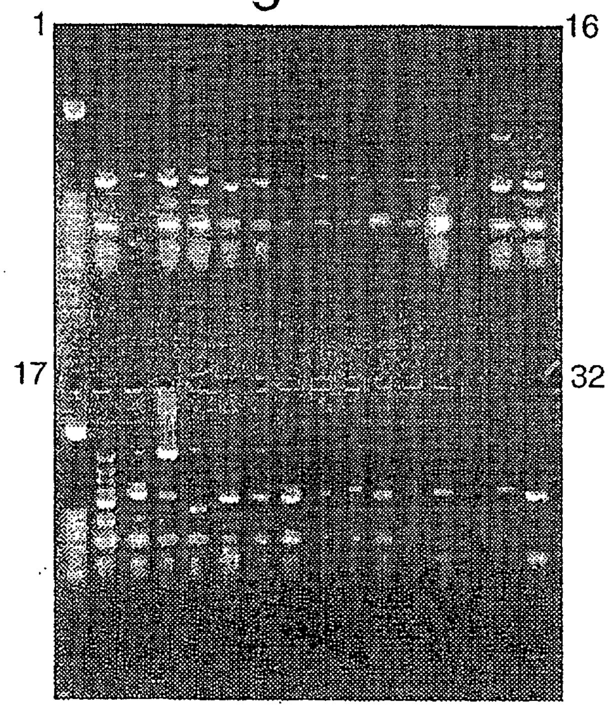
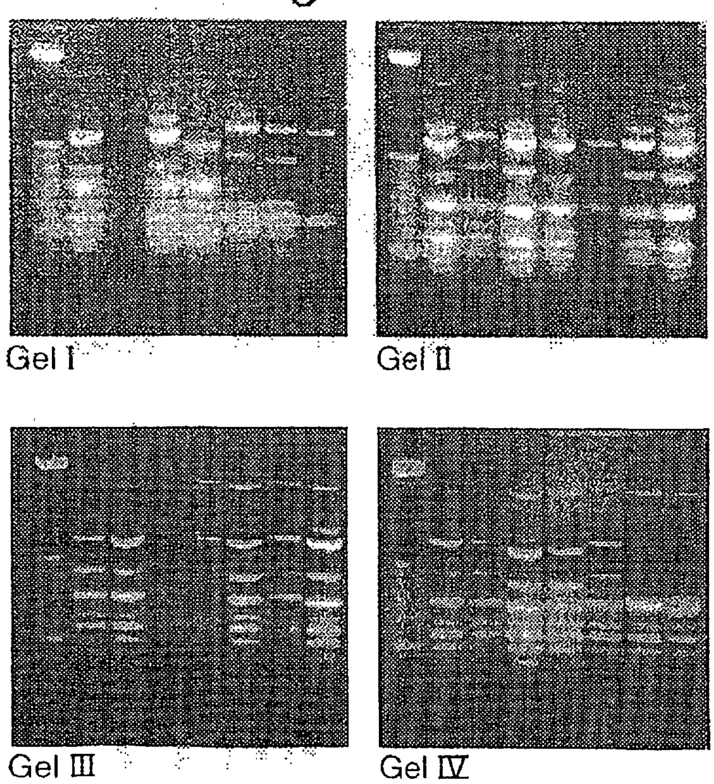


Fig.9B.



European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 03 07 6606

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
D,X	LEUNG D W ET AL: "A METHOD FOR RANDOM MUTAGENESIS OF A DEFINED DNA SEGMENT USING A MODIFIED POLYMERASE CHAIN REACTION" TECHNIQUE, vol. 1, no. 1, August 1989 (1989-08), pages 11-15, XP000614654 * the whole document * ---	1-4,15, 18-24, 37,38, 44-49	C12N15/10 C07K16/00 C12Q1/68
X	WEISBERG E P ET AL: "SIMULTANEOUS MUTAGENESIS OF MULTIPLE SITES: APPLICATION OF THE LIGASE CHAIN REACTION USING PCR PRODUCTS INSTEAD OF OLIGONUCLEOTIDES" BIOTECHNIQUES, EATON PUBLISHING, NATICK, US, vol. 15, no. 1, 1 July 1993 (1993-07-01), pages 68-70,72-74,, XP000385832 ISSN: 0736-6205 * the whole document * ---	1-4, 22-28, 44-49	
X	BERGER S L ET AL: "PHOENIX MUTAGENESIS: ONE-STEP REASSEMBLY OF MULTIPLY CLEAVED PLASMIDS WITH MIXTURES OF MUTANT AND WILD-TYPE FRAGMENTS" ANALYTICAL BIOCHEMISTRY, ACADEMIC PRESS, SAN DIEGO, CA, US, vol. 214, 1993, pages 571-579, XP002043107 ISSN: 0003-2697 * the whole document * ---	1,3,18, 19,44-49	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
			C12N C07K C12Q
X	WO 96 07754 A (SCRIPPS RESEARCH INST) 14 March 1996 (1996-03-14) * claims 1-37 * --- -/--	1-4, 15-20, 22-24, 26-28, 44-49	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	20 August 2003	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 03 07 6606

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	WO 94 18219 A (SCRIPPS RESEARCH INST) 18 August 1994 (1994-08-18) * claims 1-37 *	1-4, 15-20, 22-24, 26-28, 44-49	
X	EP 0 557 897 A (HOFFMANN LA ROCHE) 1 September 1993 (1993-09-01) * claims 1-27; figures 1-14 *	1-4, 15-19, 22-28, 44-49	
X	WO 94 03627 A (BRITISH TECHNOLOGY GROUP INC) 17 February 1994 (1994-02-17) * page 14, line 18 - line 22 * * page 16, line 19 - line 22 * * claims 1-8 *	1-4	
P,X	WO 97 08320 A (MORPHOSYS PROTEINOPTIMIERUNG ;KNAPPIK ACHIM (DE); PACK PETER (DE);) 6 March 1997 (1997-03-06) * claims 1-55; figure 1 *	1-4, 15-28, 44-49	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
A	WO 96 23899 A (UNIV MASSACHUSETTS MEDICAL) 8 August 1996 (1996-08-08) * claim 1 *		
A	WO 95 22625 A (AFFYMAX TECH NV ;STEMMER WILLEM P C (US); CRAMERI ANDREAS (US)) 24 August 1995 (1995-08-24) * the whole document *	1-19	
A	WO 96 17056 A (PASTEUR INSTITUT) 6 June 1996 (1996-06-06) * the whole document *	1-19	

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	20 August 2003	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 03 07 6606

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	WO 94 12632 A (UNIV LONDON ;PRODROMOU CHRISOSTOMOS (GB); PEARL LAURENCE HARRIS (G) 9 June 1994 (1994-06-09) * the whole document * ---	1-19	
A	EP 0 456 304 A (EASTMAN KODAK CO) 13 November 1991 (1991-11-13) * the whole document * ---	1-19	
A	US 5 023 171 A (HO STEFFAN N ET AL) 11 June 1991 (1991-06-11) * the whole document * -----	1-19	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	20 August 2003	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 03 07 6606

This annex lists the patent family membersrelating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-08-2003

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 9607754	A	14-03-1996	US	5667988	A	16-09-1997
			AT	236992	T	15-04-2003
			AU	706343	B2	17-06-1999
			AU	3504695	A	27-03-1996
			CA	2198899	A1	14-03-1996
			DE	69530305	D1	15-05-2003
			DK	779933	T3	28-07-2003
			EP	0779933	A1	25-06-1997
			JP	10504970	T	19-05-1998
			WO	9607754	A1	14-03-1996
			US	6096551	A	01-08-2000
WO 9418219	A	18-08-1994	AU	6132994	A	29-08-1994
			AU	6170394	A	29-08-1994
			AU	6235294	A	29-08-1994
			US	5679548	A	21-10-1997
			WO	9418219	A1	18-08-1994
			WO	9418220	A1	18-08-1994
			WO	9418221	A1	18-08-1994
			US	5667988	A	16-09-1997
			US	6096551	A	01-08-2000
EP 0557897	A	01-09-1993	US	5395750	A	07-03-1995
			AU	667506	B2	28-03-1996
			AU	3379193	A	02-09-1993
			CA	2089966	A1	29-08-1993
			EP	0557897	A1	01-09-1993
			JP	6121696	A	06-05-1994
			NZ	245970	A	27-04-1995
WO 9403627	A	17-02-1994	AU	4782693	A	03-03-1994
			WO	9403627	A1	17-02-1994
WO 9708320	A	06-03-1997	AT	219517	T	15-07-2002
			AU	725609	B2	12-10-2000
			AU	6874596	A	19-03-1997
			CA	2229043	A1	06-03-1997
			DE	69621940	D1	25-07-2002
			DE	69621940	T2	16-01-2003
			DK	859841	T3	09-09-2002
			WO	9708320	A1	06-03-1997
			EP	0859841	A1	26-08-1998
			ES	2176484	T3	01-12-2002
			JP	2001519643	T	23-10-2001
			PT	859841	T	29-11-2002
			US	6300064	B1	09-10-2001

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 03 07 6606

This annex lists the patent family membersrelating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-08-2003

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 9623899	A	08-08-1996	AU	3204895	A	21-08-1996
			WO	9623899	A1	08-08-1996
WO 9522625	A	24-08-1995	US	5605793	A	25-02-1997
			AT	216722	T	15-05-2002
			AT	211761	T	15-01-2002
			AU	703264	B2	25-03-1999
			AU	2971495	A	04-09-1995
			CA	2182393	A1	24-08-1995
			CN	1145641	A	19-03-1997
			DE	1094108	T1	17-10-2002
			DE	69525084	D1	28-02-2002
			DE	69525084	T2	11-07-2002
			DE	69526497	D1	29-05-2002
			DE	69526497	T2	21-11-2002
			DE	752008	T1	05-04-2001
			DE	934999	T1	05-04-2001
			DK	752008	T3	17-06-2002
			DK	934999	T3	18-02-2002
			EP	1094108	A2	25-04-2001
			EP	1205547	A2	15-05-2002
			EP	0752008	A1	08-01-1997
			EP	0934999	A1	11-08-1999
			ES	2176324	T3	01-12-2002
			ES	2165652	T3	16-03-2002
			JP	10500561	T	20-01-1998
			JP	3393848	B2	07-04-2003
			JP	2001057893	A	06-03-2001
			JP	2003033180	A	04-02-2003
			JP	2003174881	A	24-06-2003
			RU	2157851	C2	20-10-2000
			WO	9522625	A1	24-08-1995
			US	6180406	B1	30-01-2001
			US	6132970	A	17-10-2000
			US	6287861	B1	11-09-2001
			US	6291242	B1	18-09-2001
			US	2002025517	A1	28-02-2002
			US	6420175	B1	16-07-2002
			US	6277638	B1	21-08-2001
			US	6573098	B1	03-06-2003
			US	6323030	B1	27-11-2001
			US	6413774	B1	02-07-2002
			US	2003082611	A1	01-05-2003
			US	6319713	B1	20-11-2001
			US	6355484	B1	12-03-2002
			US	6309883	B1	30-10-2001

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 03 07 6606

This annex lists the patent family membersrelating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

20-08-2003

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9522625 A		US 6297053 B1	02-10-2001
		US 5830721 A	03-11-1998
		US 2002051976 A1	02-05-2002
		US 5811238 A	22-09-1998
		US 6372497 B1	16-04-2002
		US 6344356 B1	05-02-2002
		US 6395547 B1	28-05-2002
WO 9617056 A	06-06-1996	EP 0714980 A1	05-06-1996
		WO 9617056 A1	06-06-1996
		US 5843730 A	01-12-1998
WO 9412632 A	09-06-1994	WO 9412632 A1	09-06-1994
EP 0456304 A	13-11-1991	US 5387505 A	07-02-1995
		AT 130048 T	15-11-1995
		CA 2039222 A1	05-11-1991
		DE 69114353 D1	14-12-1995
		DE 69114353 T2	18-04-1996
		EP 0456304 A1	13-11-1991
		FI 912141 A	05-11-1991
		HK 38296 A	15-03-1996
		IE 911508 A1	06-11-1991
		JP 4228076 A	18-08-1992
US 5023171 A	11-06-1991	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82